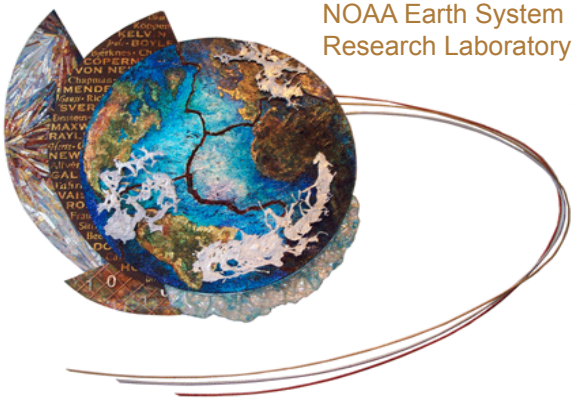


NOAA Earth System Research Laboratory

Reforecasting: status and plans

Tom Galarneau[1], Tom Hamill[2], and Jeff Whitaker[2]

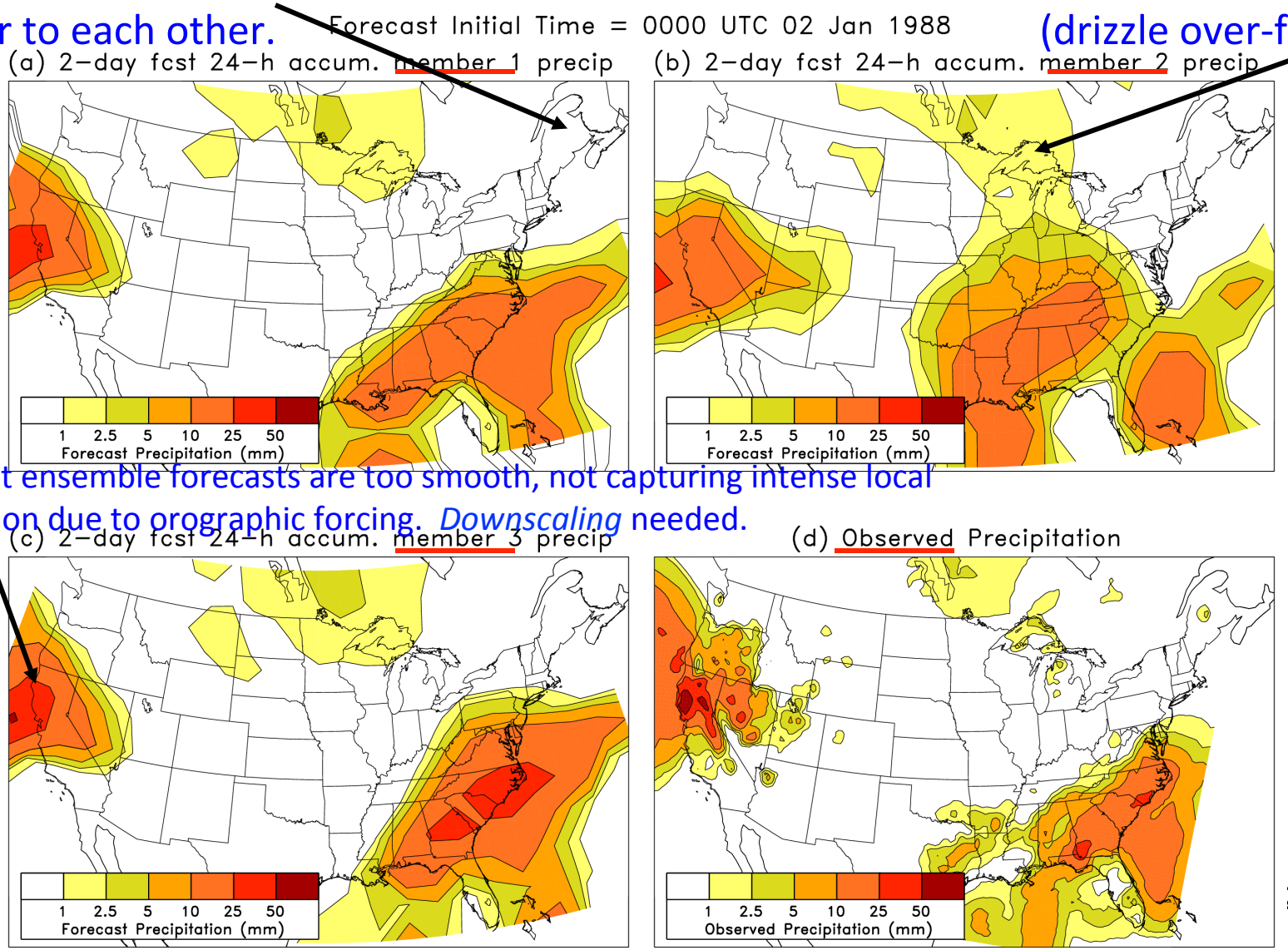
HMT West Annual Meeting, 7 Oct 2010

[1]CIRES and [2]NOAA/ESRL/PSD

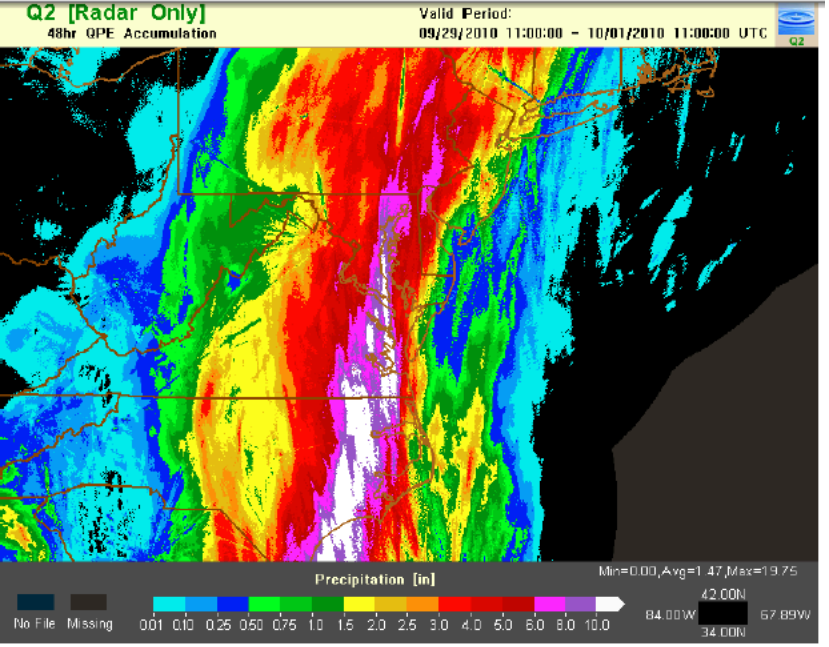
Imperfect ensembles: a few problems we'd like to correct through post-processing.

(2) ensemble members too similar to each other.

(1) conditional bias (drizzle over-forecast)



(3) Current ensemble forecasts are too smooth, not capturing intense local precipitation due to orographic forcing. *Downscaling* needed.



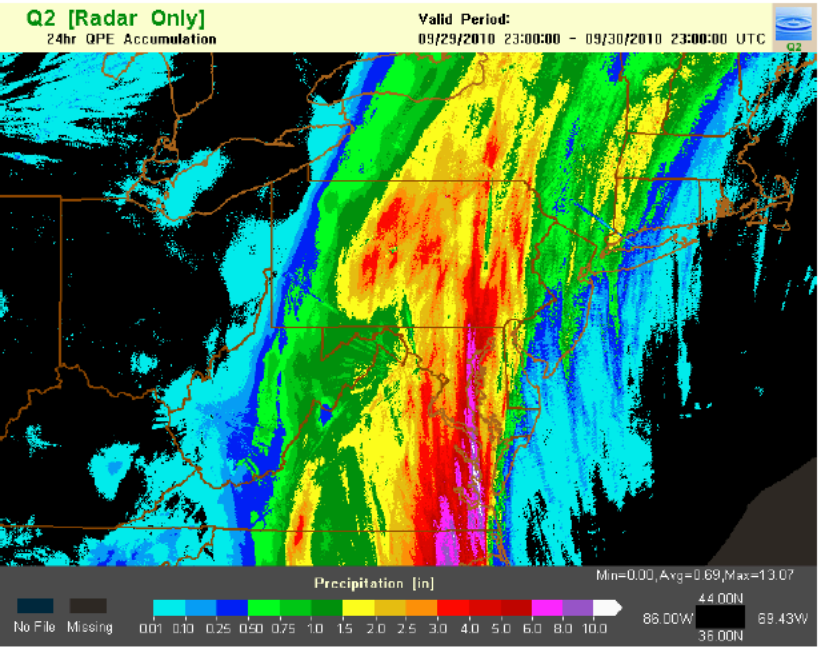


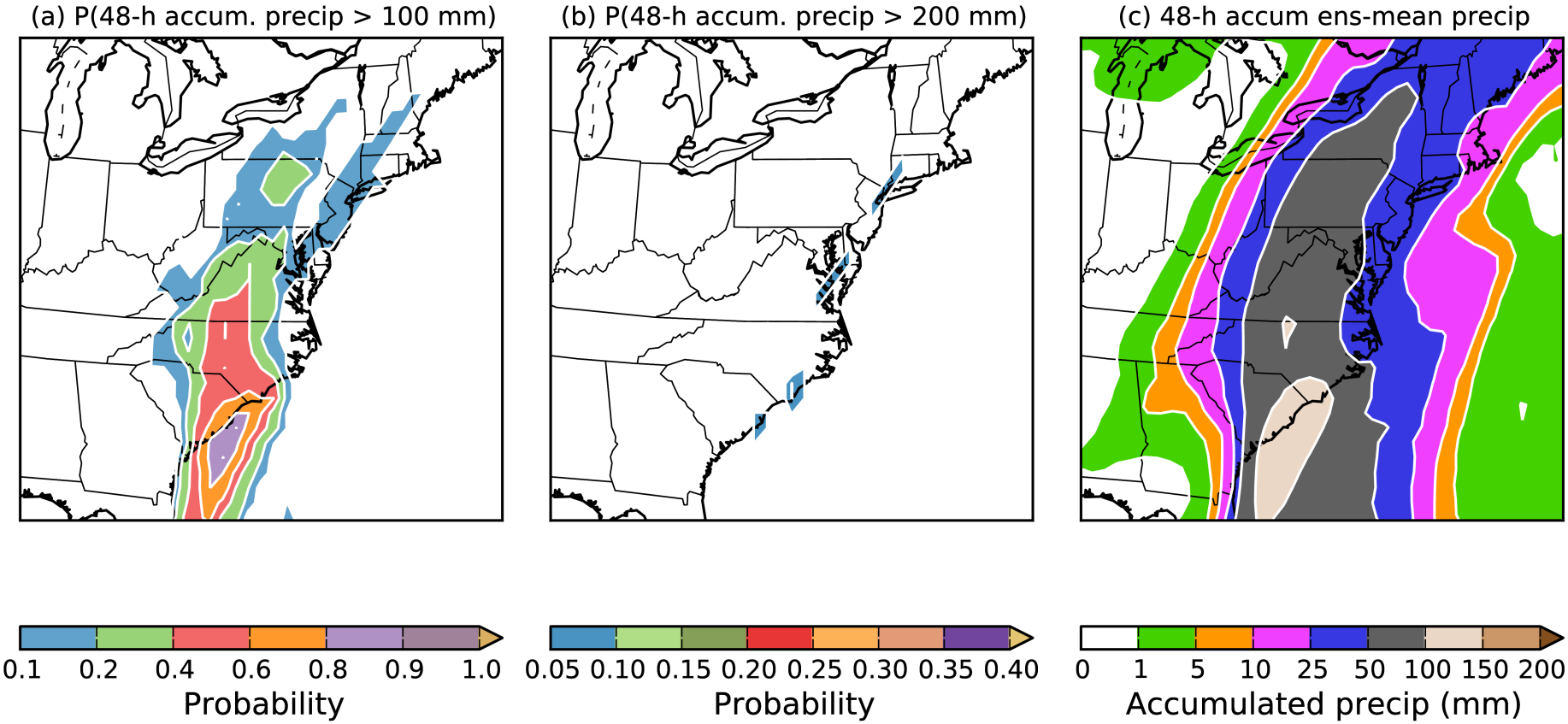
Figure 1. Q2 precipitation analysis over the Mid-Atlantic region. Upper panel is a 2-day total ending at 1100 UTC 1 October 2010 and the lower panels of the last 24 hours ending 2300 UTC 30 September 2010.

Preliminary precipitation analysis, c/o Rich Grumm, NWS/WFO, State College PA

Massive and widespread East-coast rain event, linked in part to moisture advected ahead of remnants of tropical storm Nicole.

Probability and ensemble mean, GFS/EnKF

T254 GFS/EnKF 24-72 hour forecast from 2010092812



Axis of heavy precipitation progged too far west, but still signal of major event.

Two ways NOAA is pushing forward on PQPF research

(1) High-resolution ensembles with explicitly resolved convection – "dynamical downscaling"

(2) Post-processing of lower-resolution ensembles using reforecasts – "statistical downscaling"

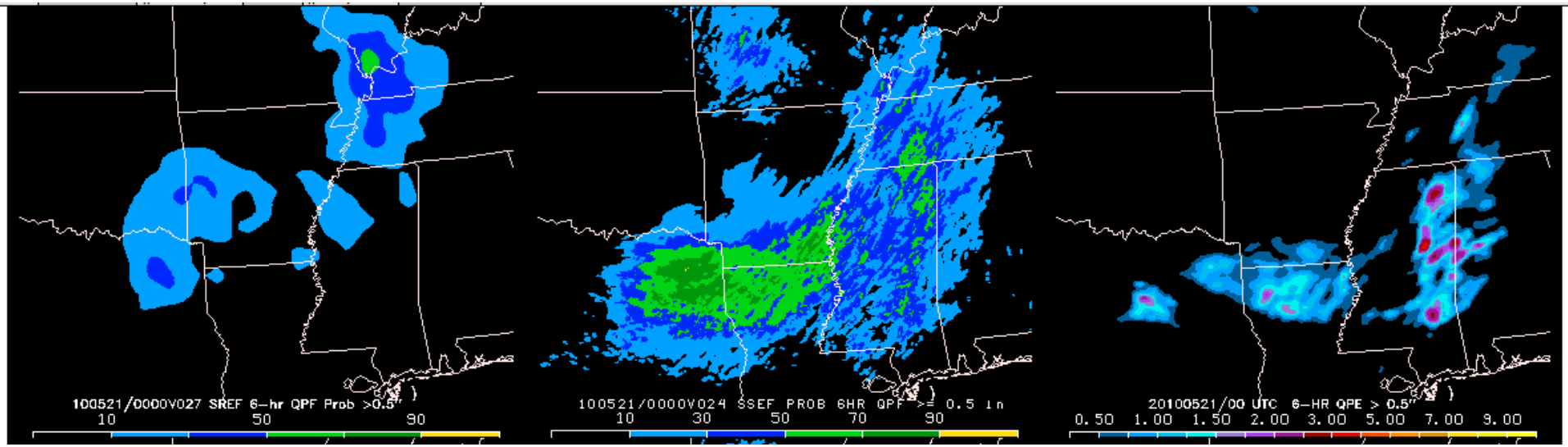
Need to understand merits and limitations of each

Potential value of storm-scale ensemble forecasts (SSEF)

An example from NSSL-SPC Hazardous Weather Test Bed, forecast initialized 20 May 2010

http://tinyurl.com/2ftbvgs

SREF P > 0.5” 4-km SSEF P > 0.5 “ Verification

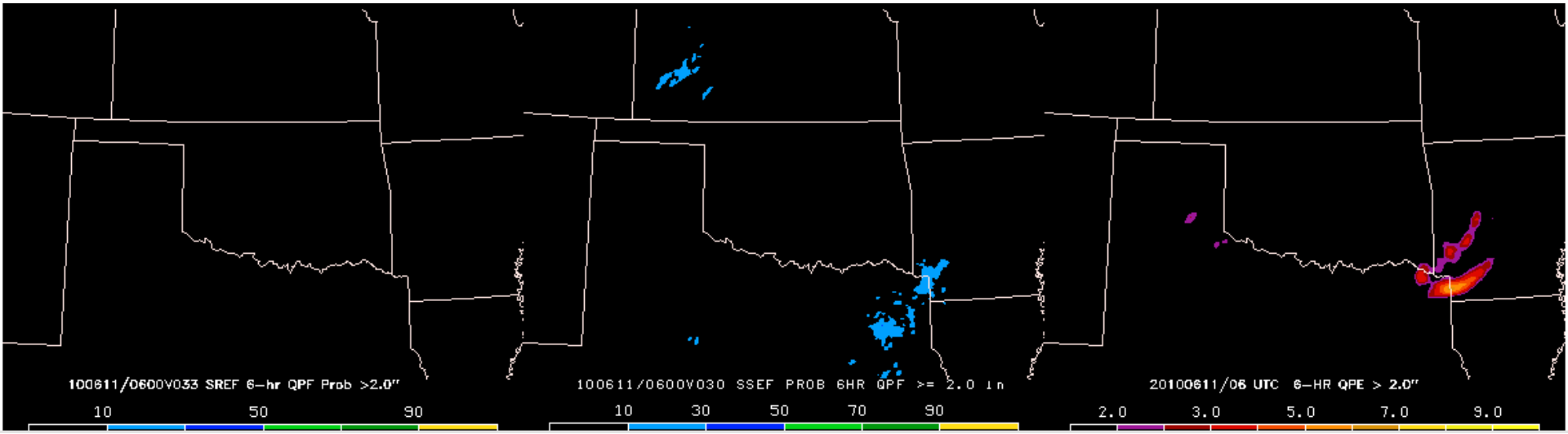


With warm-season QPF, coarse resolution and parameterized convection of SREF clearly inferior to the 4-km, resolved convection in SSEF.

A counterpoint: Arkansas floods

An example from NSSL-SPC Hazardous Weather Test Bed, forecast initialized 10 June 2010; many deaths in campground from flash flood; http://tinyurl.com/34568hp

SREF P > 2.0” 4-km SSEF P > 2.0 “ Verification (radar QPE)



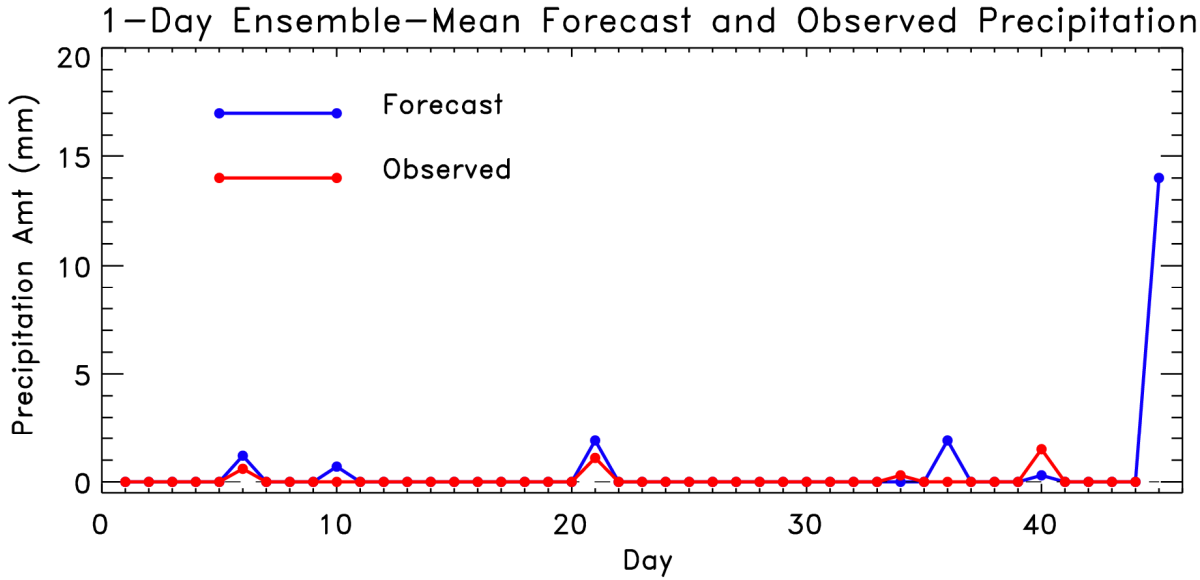
A less than 30% probability of > 2 inches rainfall from SSEF, while better than SREF, probably does not set off alarm bells in forecasters’ heads.

Reforecasts as an alternative

- What is a reforecast?
  - A data set of past forecasts using the same data assimilation system and forecast model that is used to produce the real-time forecast.
  - Ideally, multi-decadal.
- Why?
  - To assist the forecaster in interpreting model guidance, to "back out" the systematic errors of the model.

Why multi-decadal?

- To be able to correct systematic errors in numerical weather predictions, even for rare events.



Suppose today's numerical forecast is quite wet (and this is a climatologically dry region). A short time series of past forecasts and observations is not enough to provide cases similar to today's forecast, which are needed in order to make statistical corrections. The longer into the past you have a series of reforecasts, the easier it is to find samples similar to today's weather event, and the more informed statistical corrections one can make.

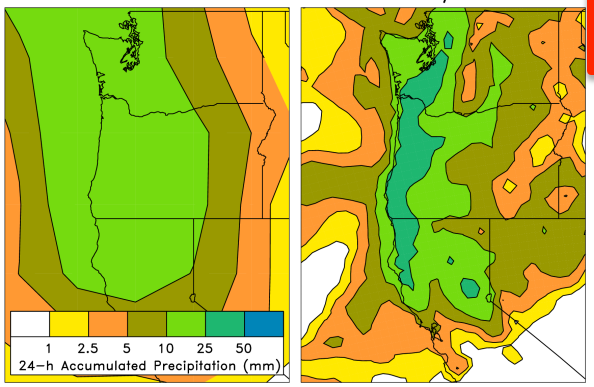
1st-generation reforecast (created in 2003 & maintained since)

- Used 1998 version of the NCEP GFS model
- Produced 15-member ensemble forecasts, once daily, to 15-days lead, *from 1979–present*
- Archived a few variables on a 2.5-degree grid (several TB of data, hosted locally at ESRL/PSD).
- Used semi-operationally at CPC, HPC, but not fully integrated with NWS operations.

An example of a statistical correction technique using those reforecasts

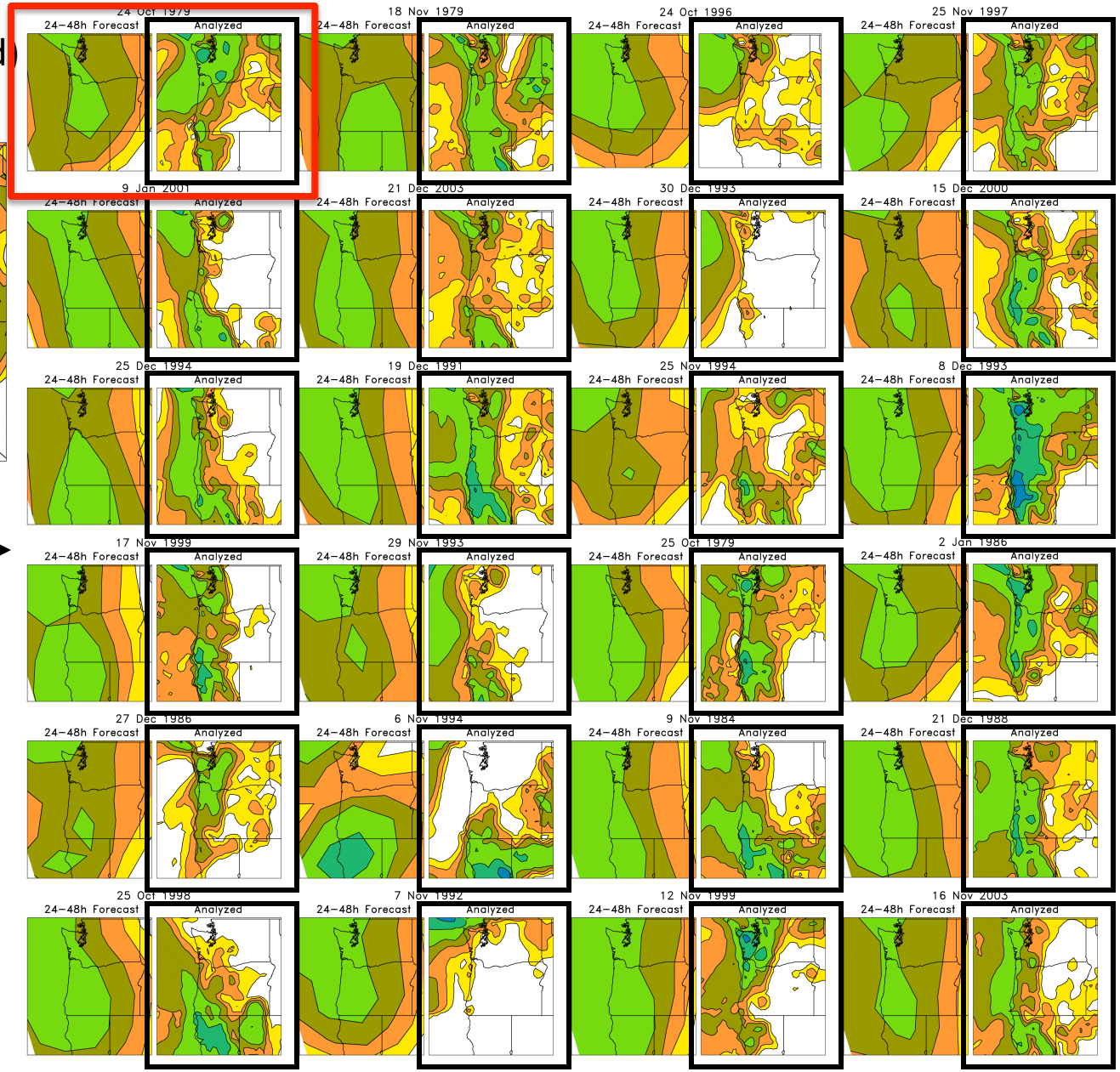
Today's forecast (& observed)

26 Nov 2005



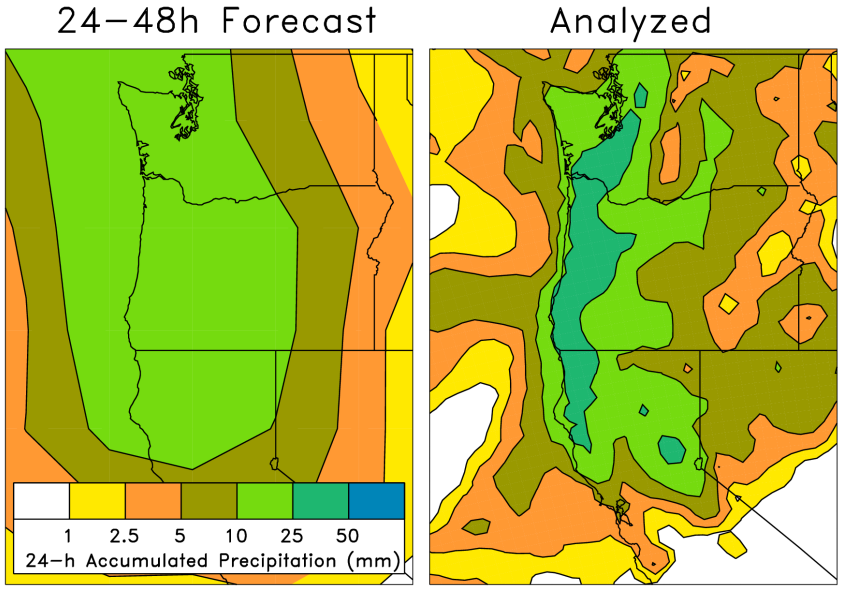
For each pair (e.g. red box), on the left are old forecasts that are somewhat similar to this day's ensemble-mean forecast. The boxed data on the right, the analyzed precipitation for the same dates as the chosen analog forecasts, can be used to statistically adjust and downscale the forecast.

Analog approaches like this may be particularly useful for hydrologic ensemble applications, where an ensemble of weather realizations is needed as inputs to a hydrologic ensemble streamflow system.

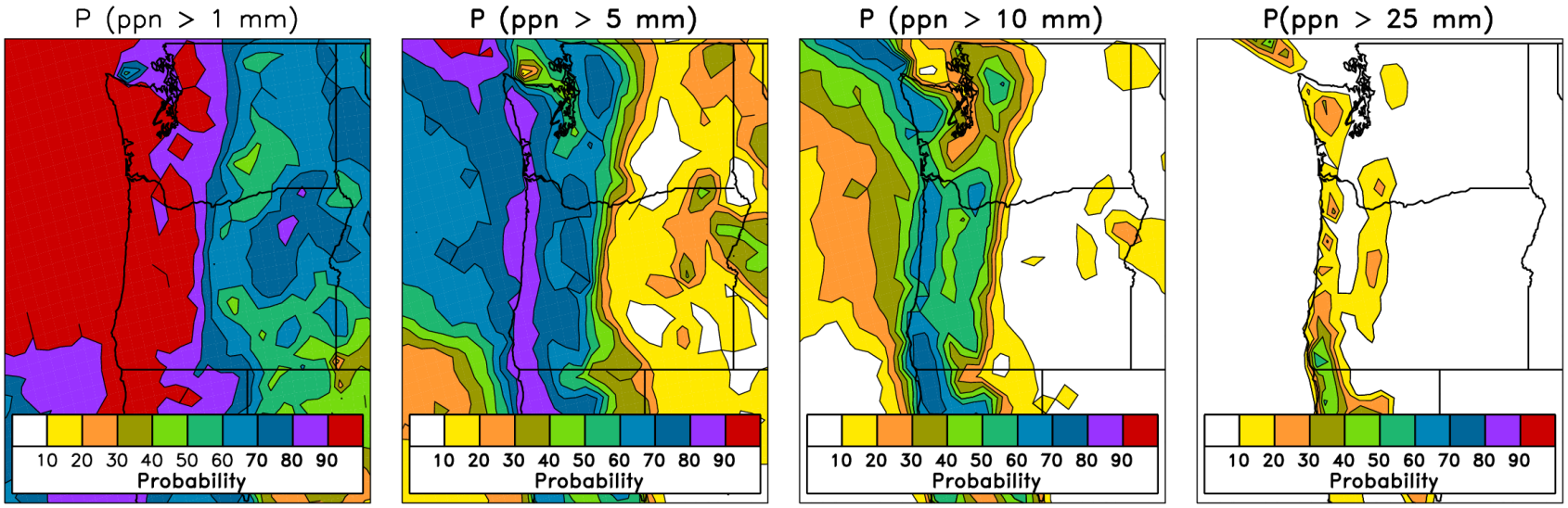


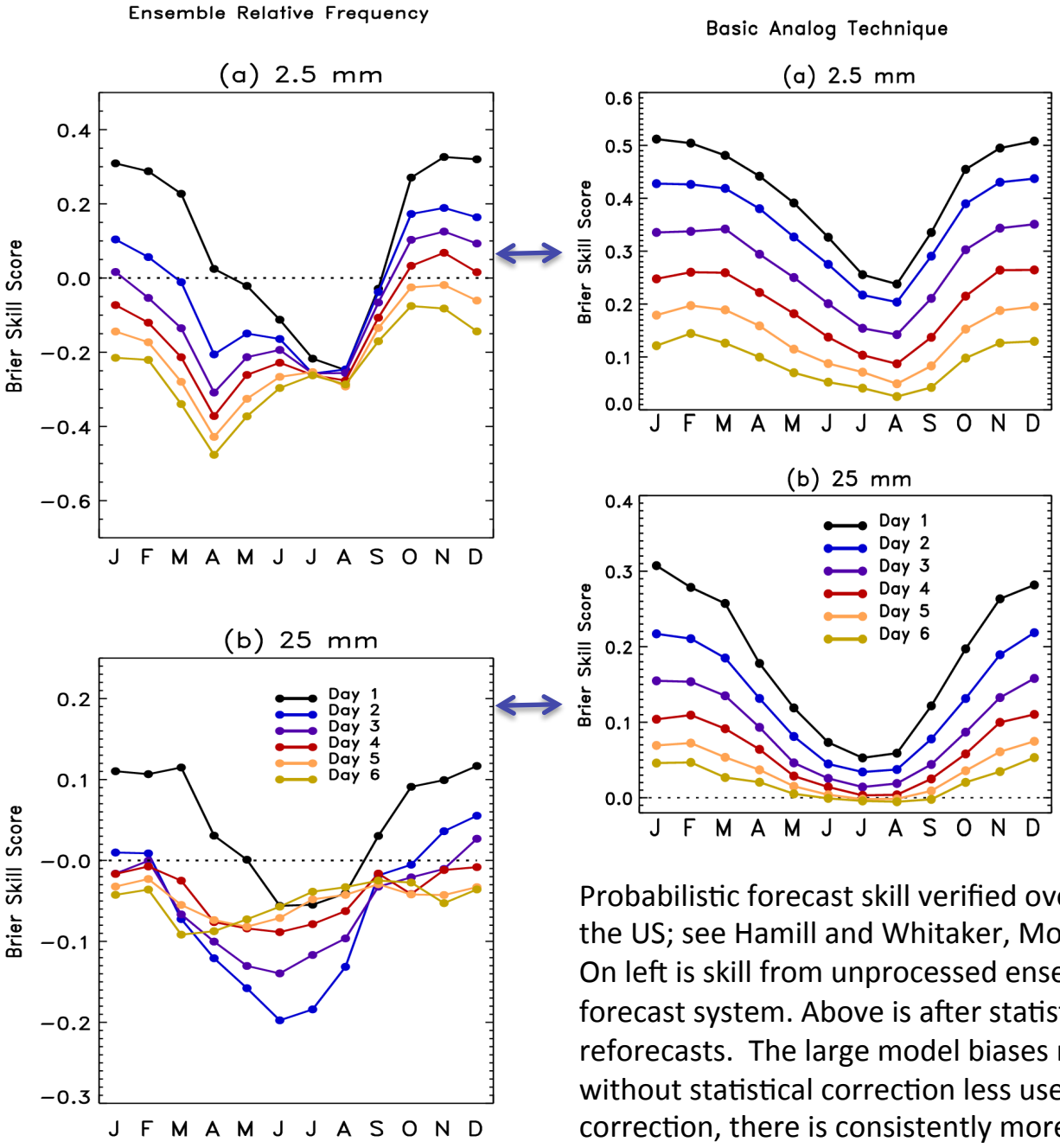
Downscaled analog probability forecasts

26 Nov 2005



Statistically adjusted analog forecast probabilities





Does this improve the forecast?

Yes, statistical correction techniques like these have produced dramatically improved forecast guidance.

Probabilistic forecast skill verified over 25 years of forecasts across the US; see Hamill and Whitaker, Monthly Weather Review, 2006. On left is skill from unprocessed ensemble from 1998 GFS ensemble forecast system. Above is after statistical post-processing using reforecasts. The large model biases make the ensemble forecasts without statistical correction less useful than climatology. After correction, there is consistently more skill than climatology.

What's the magical trick here?

- The magic IS NOT in the particular statistical technique (logistic regression, neural net, others would have done just as well as, or better than analogs).

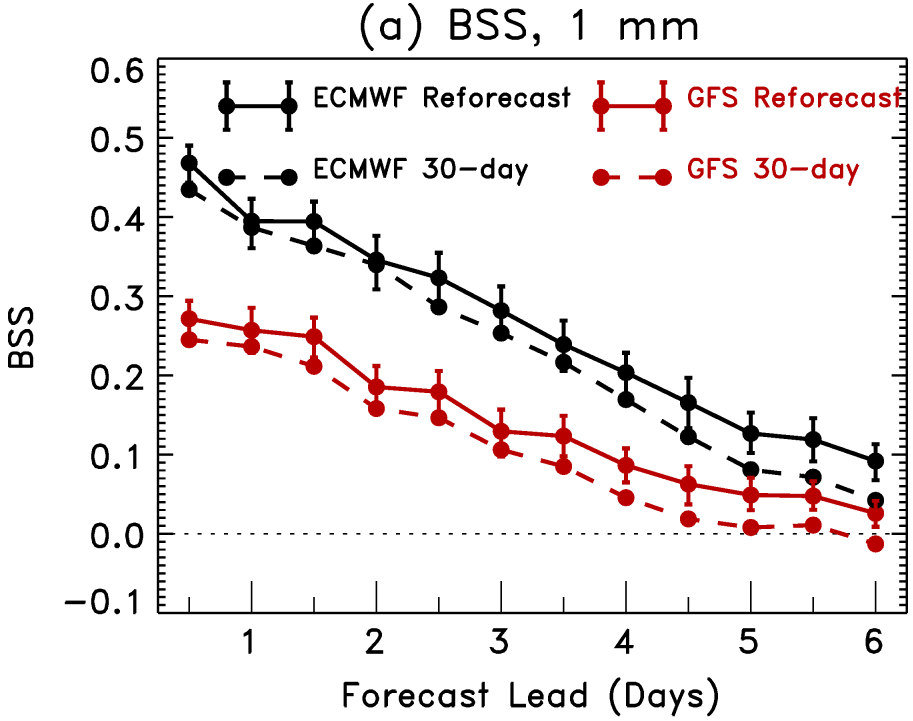
- The magic IS in the large training data set, the ability to find past cases that closely resemble today's weather and then use the distribution of observed weather events.

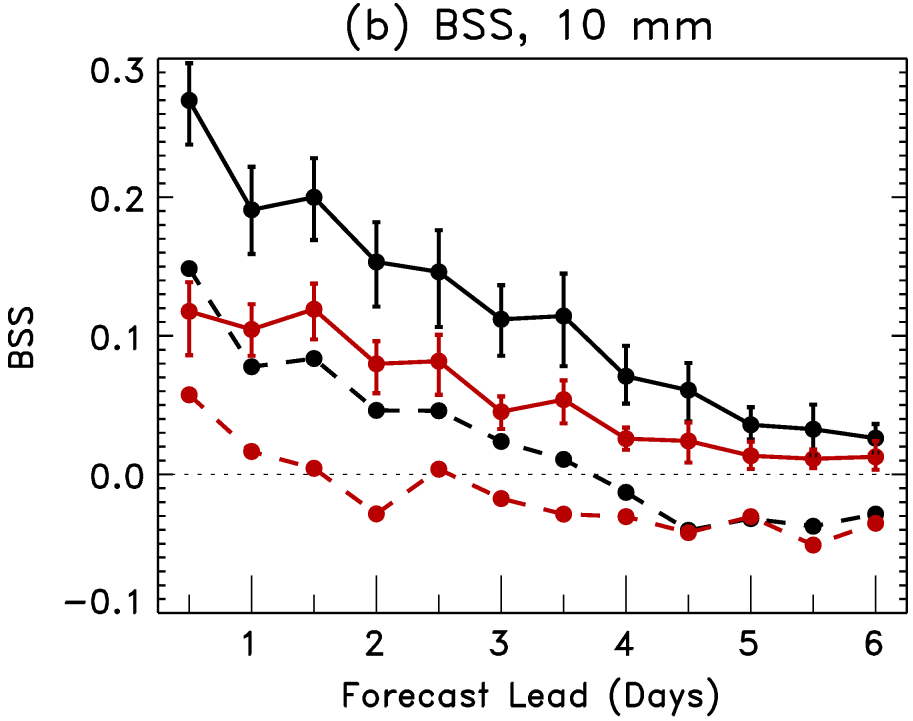
Known strengths and weaknesses of reforecasts for PQPF calibration

- Reforecasts can correct for repeatable systematic errors, e.g., big rain shield over central CA from coarse-resolution model, past events from this model suggest to enhance rainfall on mountain upslopes, dry it out in San Joaquin valley.

- Reforecasts cannot correct for unrepeatable errors. If precipitation blob is mislocated by 500 km, and it wasn't mislocated by similar amounts in past forecasts, reforecasts can't correct. Then you just need a better forecast model.

If we were reforecasting with a better model, we'd make even better PQPF's





Based on our results, ECMWF in 2005 produced a 15-member reforecast, once weekly, over a 20-year period. Their reforecast calibrated precipitation forecasts were more skillful than ours. If we reforecast using a better model, we'll be able to produce better post-processed forecasts, too.

2005 ECMWF vs. 1998 GFS

A 2nd-generation GFS reforecast

- We submitted a proposal to DOE to use their supercomputers to generate a new reforecast data set with a modern reanalysis initial condition and a modern version of the GFS model and ensemble system.
- We won the award in June 2010.
- Now we need to store the data and permit a wide variety of users to access the reforecast data that they need.
- Hope to have this dataset more broadly used throughout NWS; at NCEP, MDL, CPC, HPC, RFCs, WFOs.

What we have been granted by DOE, and what's expected of us

- 14.5 M CPU hours (CPUh) on "Franklin" supercomputer at Lawrence Berkeley lab, all to be used before June 2011.
  - some of those cycles (10%) needed for generating initial conditions, too.
- We are expected to make this data set available to the community by mid-2011
  - Will continue to generate PQPF products.

Operating principle

- Reforecasts will be computed with a (smaller-ensemble) version of the GEFS (global ensemble forecast system) that will be operational in 2011.
- We hope that GEFS will remain in this configuration, or will be changed only slightly, for several years thereafter.
- Once GEFS changes, either EMC or ESRL will continue to run the reforecast version until a next-generation reforecast is in place.

Some more details on next-generation reforecast

- Anticipated real-time GEFS configuration at NCEP next year: T254L42 (~70 km) to 8 days, T190 thereafter. ETR initial perturbations. 21 members.
- Our 00Z reforecast configuration: full 10-member forecast every day for 30 years out to 16 days.
- Our 12Z: T254L42 ensemble to day 8, 5 members, reforecast every other day
- Save the "important" reforecast data to disk, make it readily available to community.
- Save the full model fields to tape, for occasional users such as those wishing to run retrospective limited-area forecasts.

Potential impact to WFOs, RFCs, other hydro folks

- Our existing experimental PQPF products (see web page) will be converted over to use new reforecasts.
- CPC, HPC, NCEP, MDL will potentially use for their product development in next ~3-5 years.
- Raw reforecast data (& realtime) will be available to you via web interface if you wish to test your own post-processing ideas.

Some mesoscale ensemble / reforecast questions to answer within HMT & THORPEX

- In what synoptic situations is SSEF dynamical downscaling a clearly preferable tool? In what situations is global + reforecast statistical downscaling preferable?
- Are there repeatable errors in heavy precip. events evident in reforecasts? Do they suggest a cause?
- What techniques (compositing training data over many locations) may minimize the reforecast computations, making the approach more practical?
- Can lessons from one be applied to the other? e.g., from global reforecasts can we learn how to do reforecasts for SSEF?

Conclusions

- NOAA-ESRL working on a next-generation reforecast.
- With that in place a year hence, we will
  - provide you with improved experimental PQPFs.
  - feed the data to MDL, NCEP, HPC, CPC, others for them to develop products.
  - do research to understand the merits of reforecasting vs. SSEF.
  - work with NCEP on determining how to get this concept into operations.

Extra slides

Proposed fields for “fast” archive

- “Fast” archive will be on disk, readily accessible (as opposed to full model output that may be on some slower archival system)
- Mandatory level data: Geopotential height, temperature, u, v, specific humidity at 1000, 925, 850, 700, 500, 300, 250, 200, 100, 50, and 10 hPa.
- PV (K m^2 kg^{-1} s^{-1}) on θ = 320K surface.
- Wind components, potential temperature on 2 PVU surface.
- Fixed fields saved once:
  - field capacity
  - wilting point
  - land-sea mask
  - terrain height

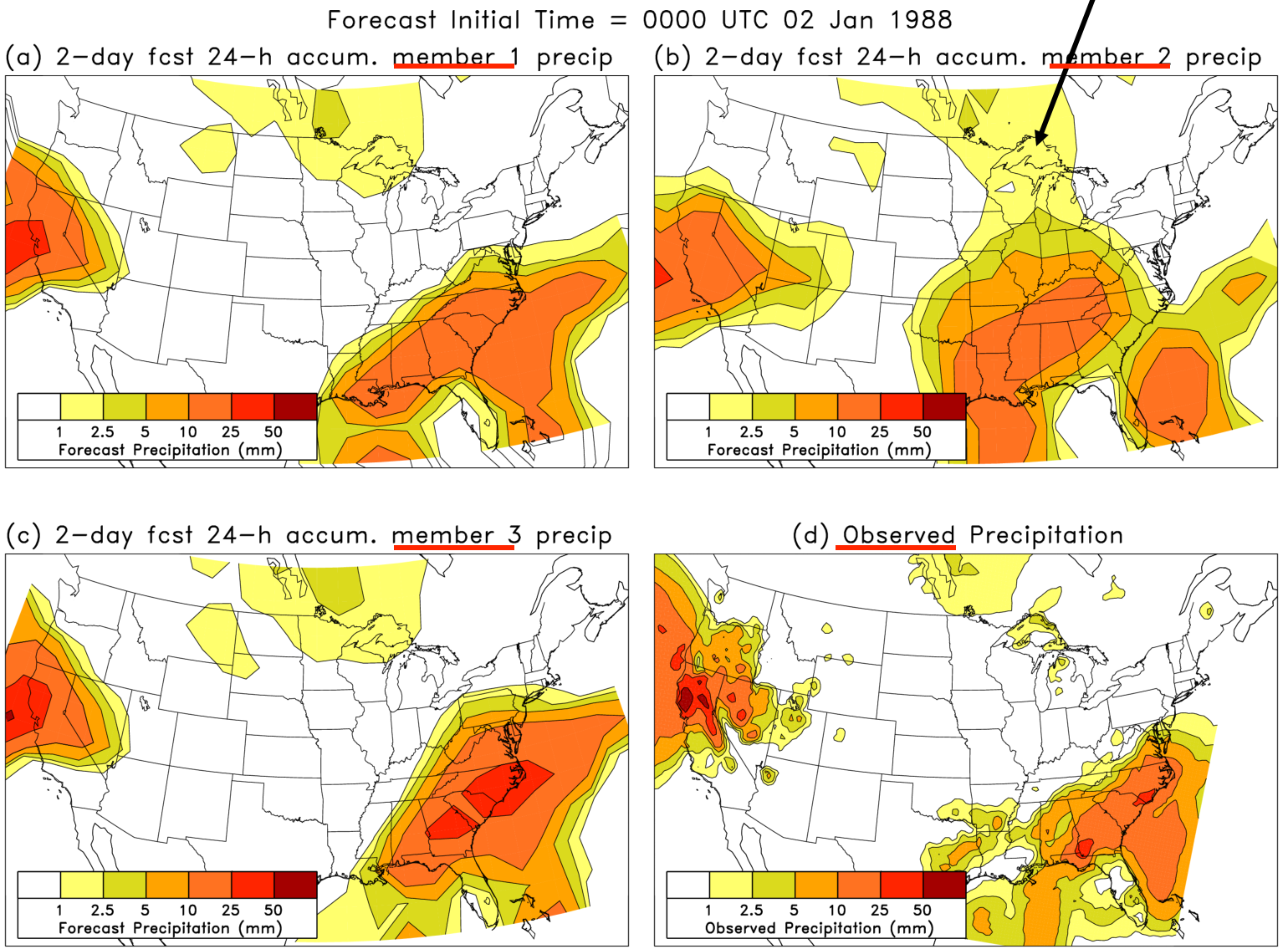
Proposed single-level fields for "fast" archive

- Surface pressure (Pa)
- Sea-level pressure (Pa)
- Surface (2-m) temperature (K)
- Skin temperature (K)
- Maximum temperature since last storage time (K)
- Minimum temperature since last storage time (K)
- Soil temperature (0-10 cm; K)
- Volumetric soil moisture content (proportion, 0-10 cm)
- Total accumulated precipitation since beginning of integration (kg/m^2)
- Precipitable water (kg/m^2, vapor only, no condensate)
- Specific humidity at 2-m AGL (kg/kg; instantaneous)
- Water equivalent of accumulated snow depth (kg/m^2)
- CAPE (J/kg)
- CIN (J/kg)
- Total cloud cover (%)
- 10-m u- and v-wind component (m/s)
- 80-m u- and v-wind component (m/s)
- Sunshine duration (min)
- Snow depth water equivalent (kg/m^2)
- Runoff
- Solid precipitation
- Liquid precipitation
- Vertical velocity (850 hPa)
- Geopotential height of surface
- Wind power (=windspeed^3 at 80 m*density)

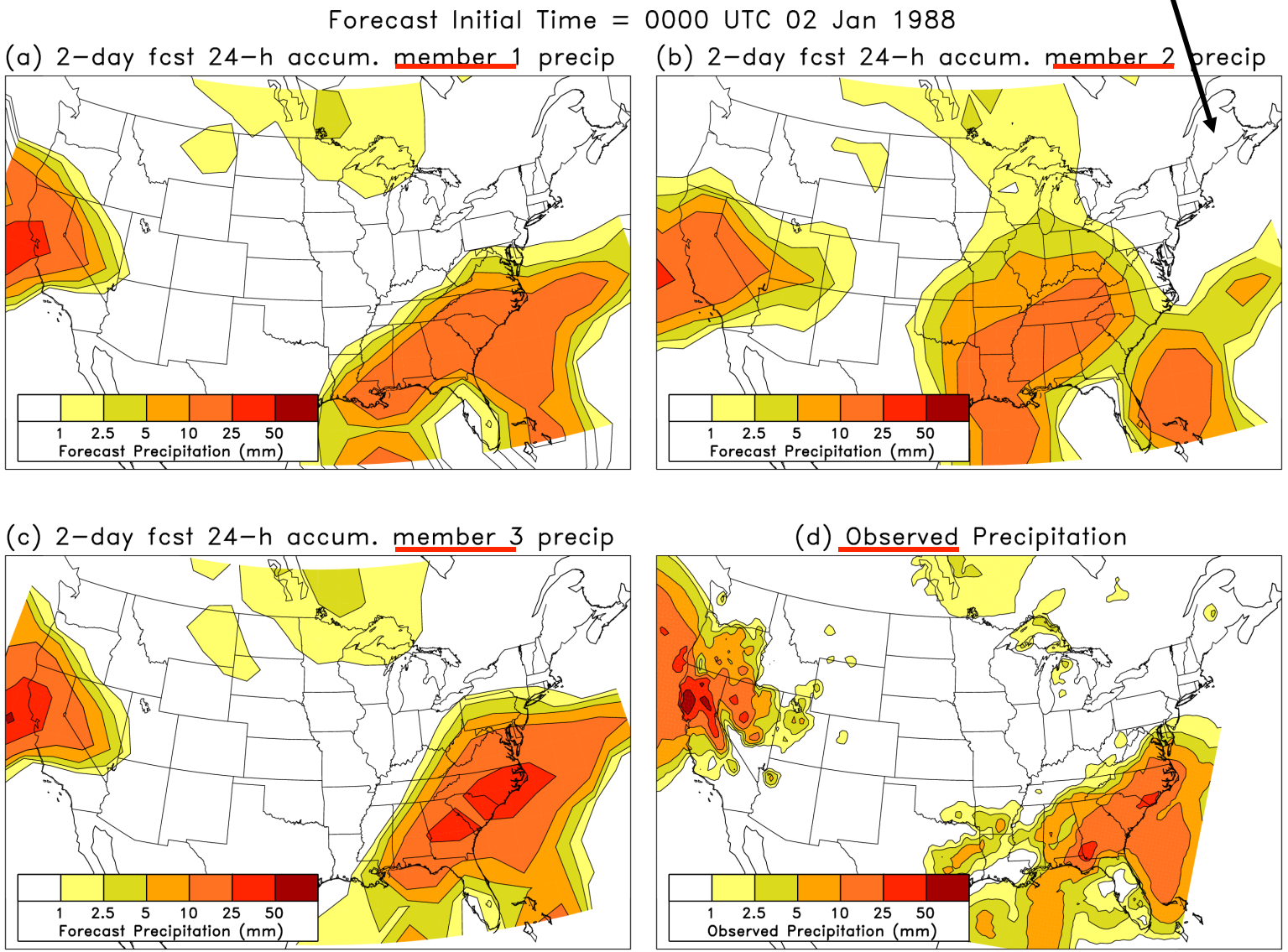
Proposed fields for “fast” archive

- Fluxes (W/m^2 ; average since last archive time)
  - sensible heat net flux at surface
  - latent heat net flux at surface
  - downward long-wave radiation flux at surface
  - upward long-wave radiation flux at surface
  - upward short-wave radiation at surface
  - downward short-wave radiation flux at surface
  - upward long-wave radiation at nominal top
  - ground heat flux.

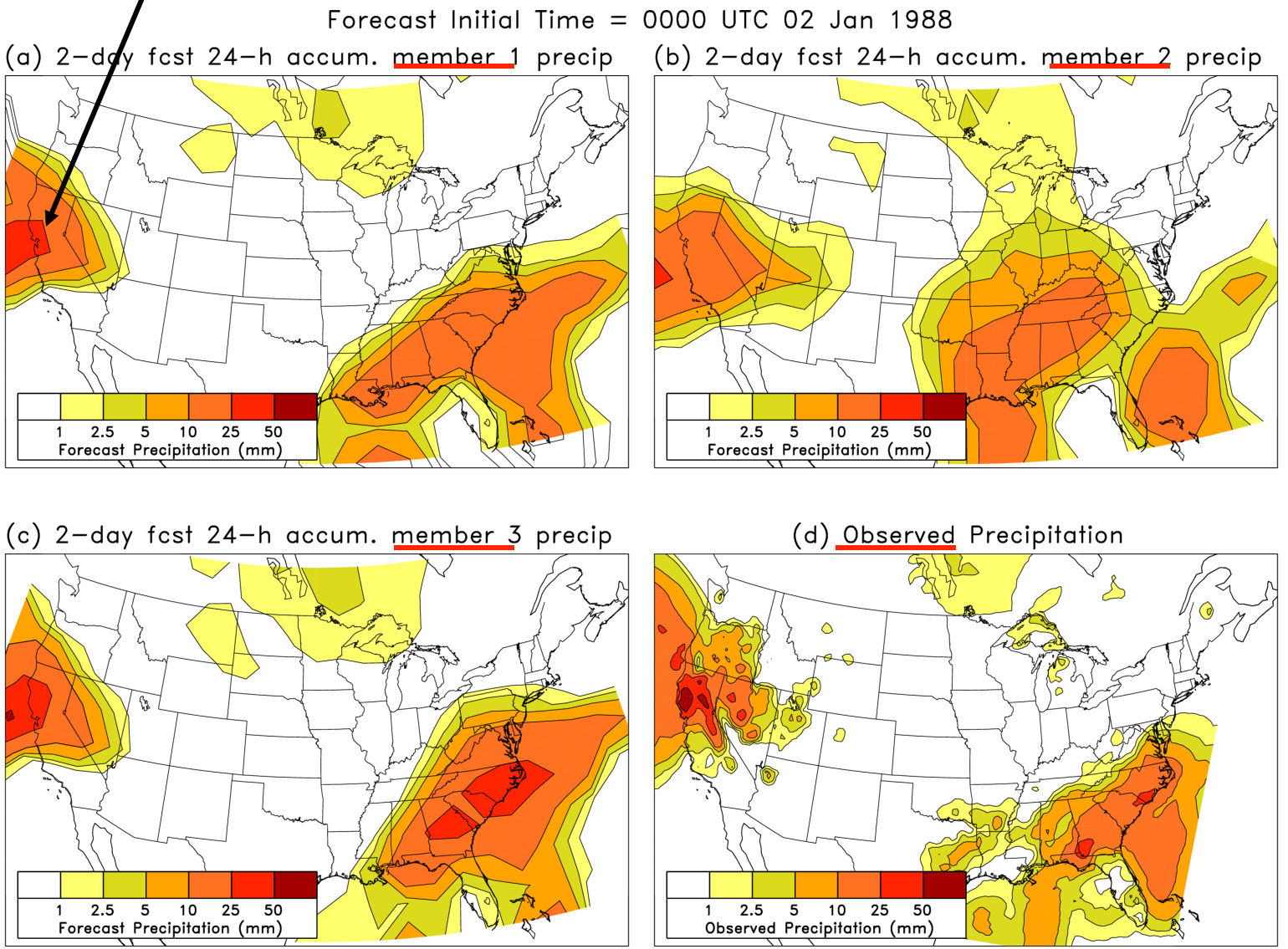
(1) conditional bias
(drizzle over-forecast)



(2) ensemble members too similar to each other.

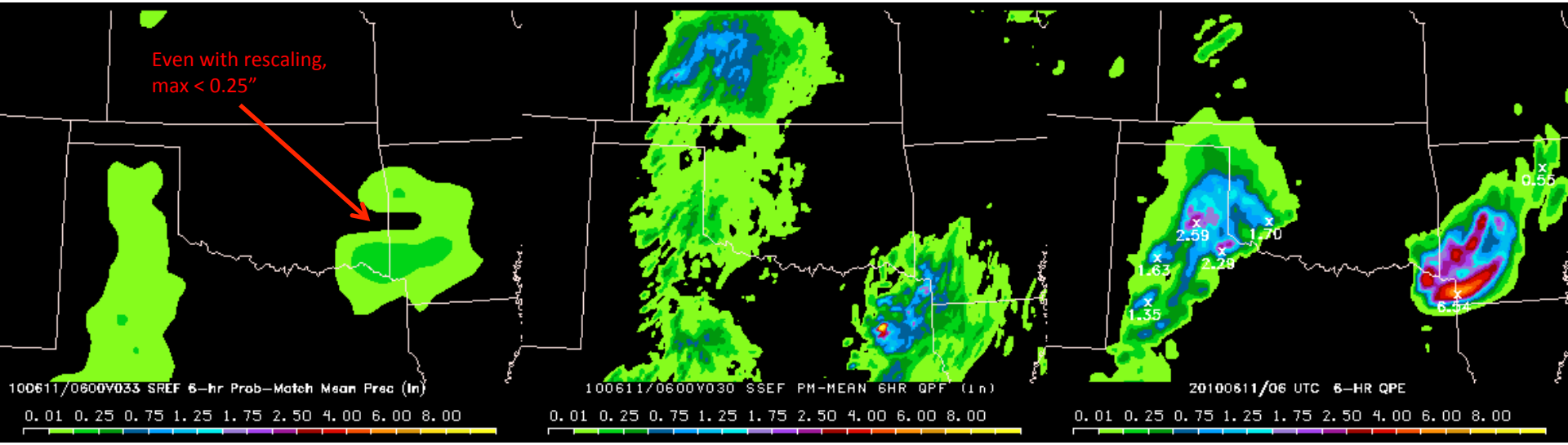


(3) Current ensemble forecasts are too smooth, not capturing intense local precipitation due to orographic forcing. *Downscaling* needed.



More from Arkansas floods

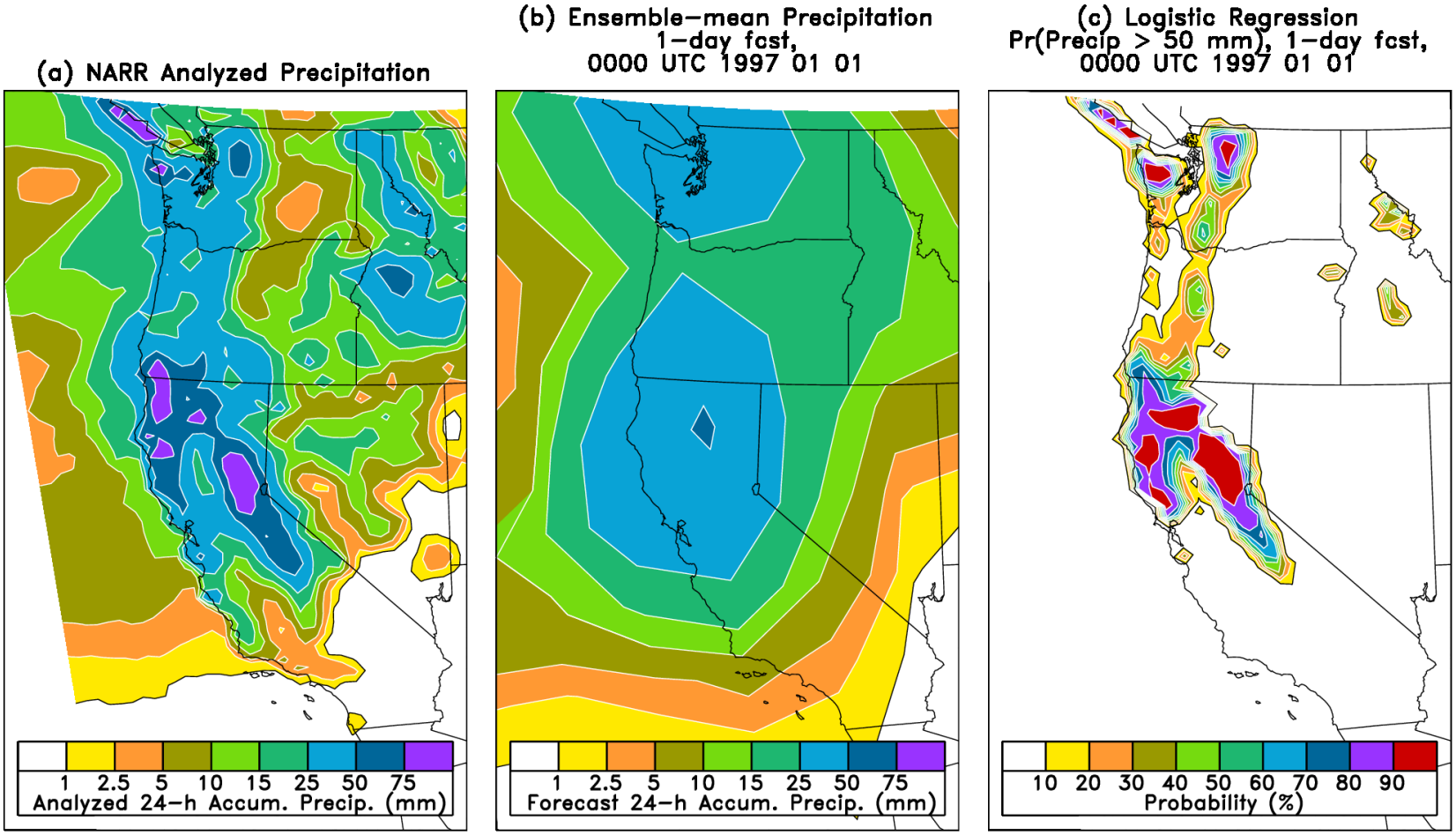
SREF P-matched mean | SSEF P-matched mean | Verification (radar QPE)



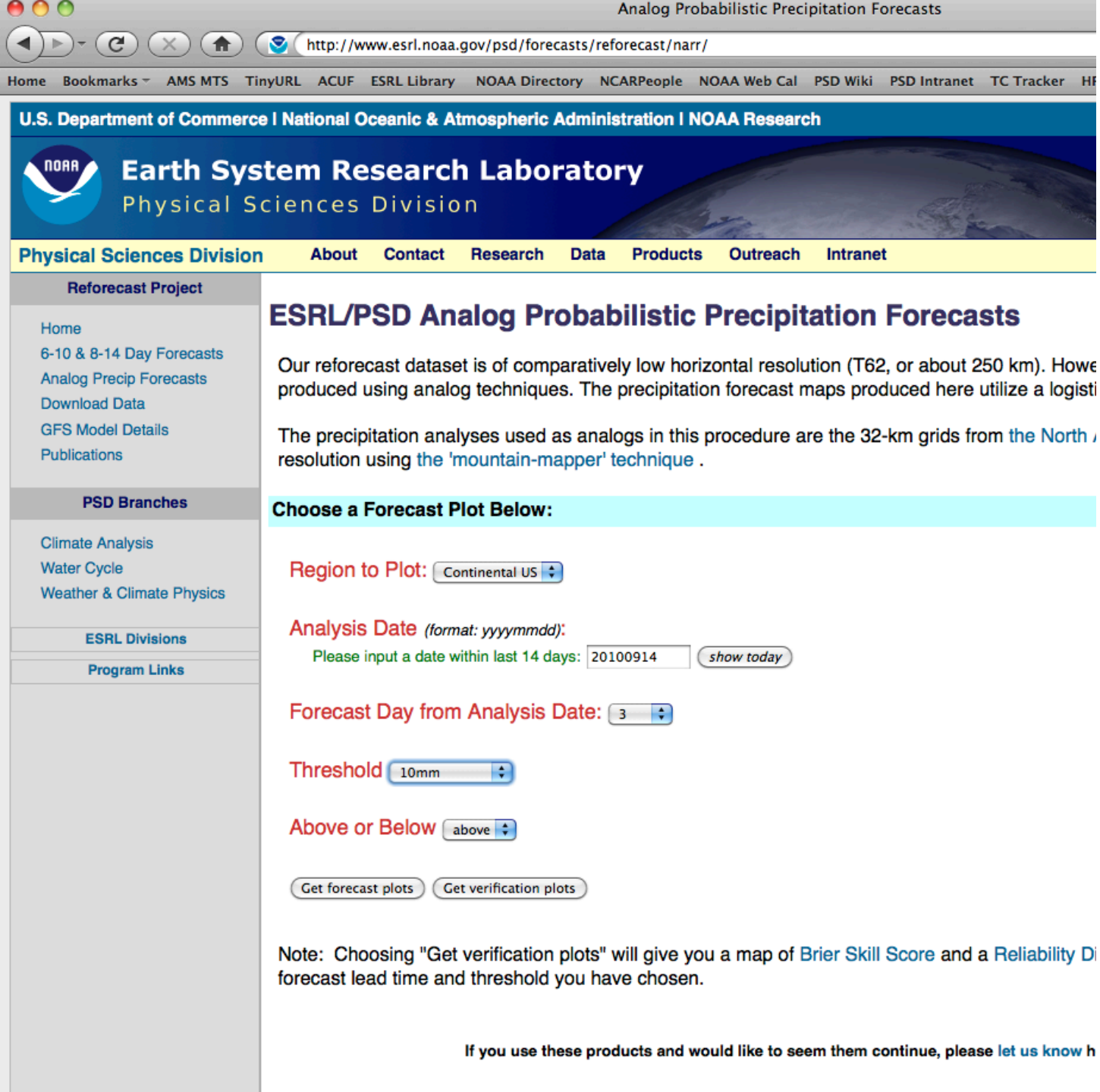
"Probability-matched mean" is a technique by Beth Ebert that rescales the ensemble-mean forecast so it has the same cumulative distribution, low-to-high, as the ensemble members. In this way its values are not smeared out like they are in the ensemble-mean forecast.

SREF terrible, SSEF still not very impressive.

Potential value of reforecast approach

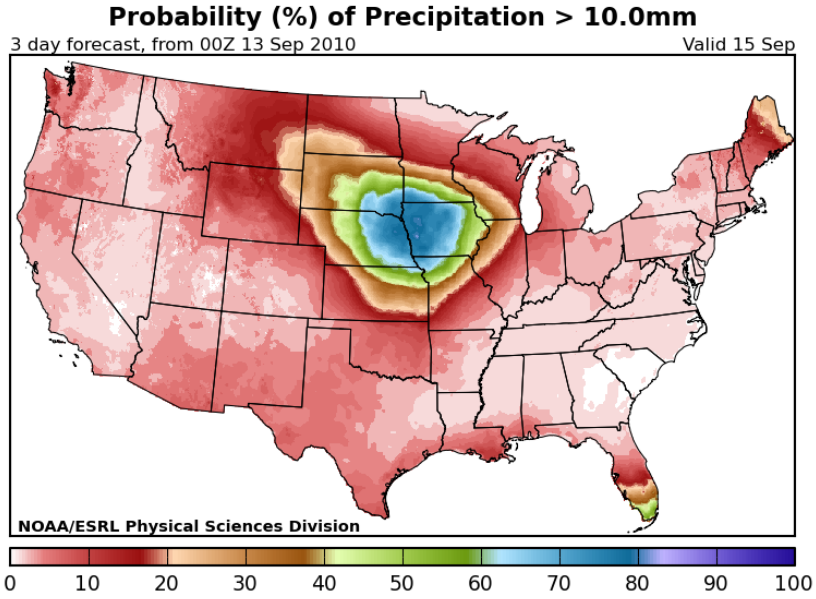


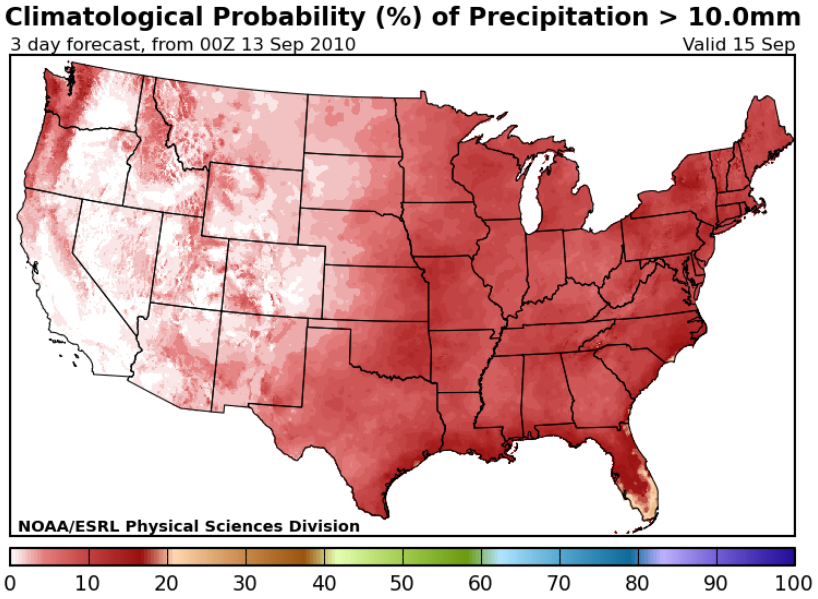
Post-processing with large training data set can permit small-scale detail to be inferred from large-scale, coarse model fields that appear very biased.



Web page to generate PQPF forecasts from 1998 GFS reforecasts

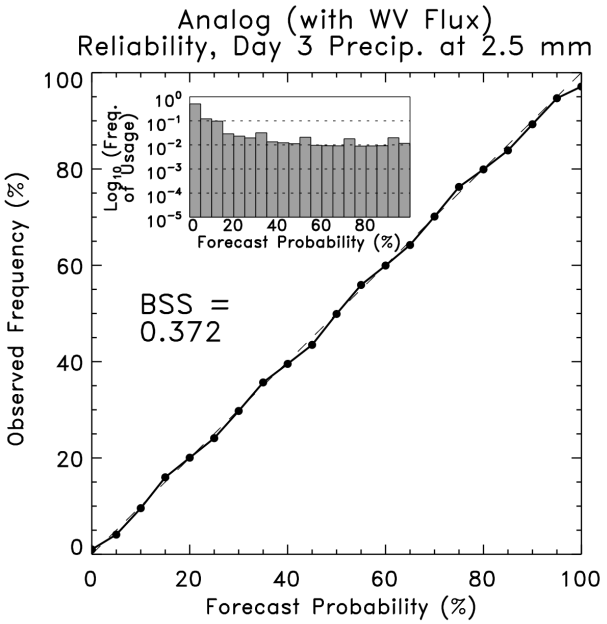
Sample plots generated



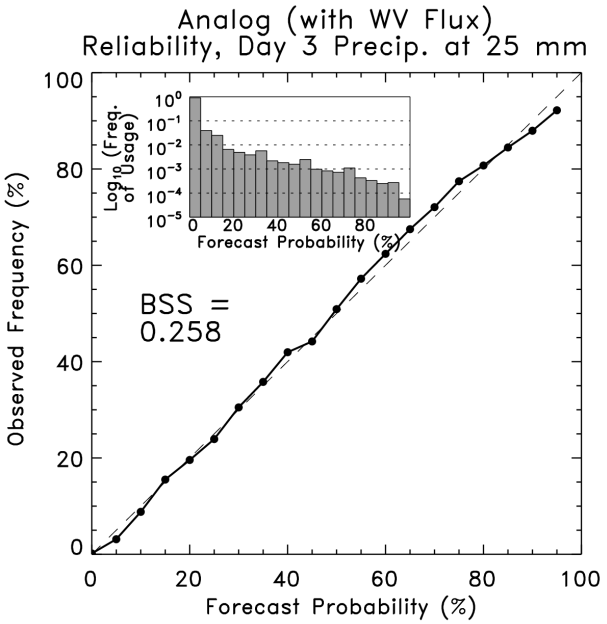


Are the forecasts reliable? Yes.

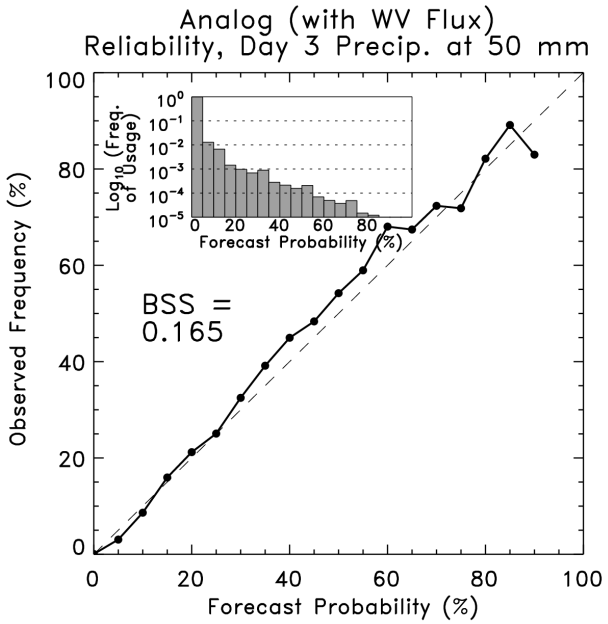
2.5 mm



25 mm



50 mm



Here verified over the West Coast over the last 25+ years, using 32-km NARR precipitation analysis

Precipitation calibration example